경제학 석사학위논문

Identification and Estimation of Additive Competing Risks Models with Unknown Transformation of Latent Failure Times

2013년 8월

서울대학교 대학원
경제학부
오 나 래

Identification and Estimation of Additive Competing Risks Models with Unknown Transformation of Latent Failure Times

지도교수 이 석 배

이 논문을 경제학석사 학위논문으로 제출함

2013년 6월

서울대학교 대학원
경제학부 경제학전공
오 나 래

오나래의 석사학위논문을 인준함

2013년 6월

위 원 장	김 재 영	(인)
부 위 원 장	이 석 배	(인)
위 원	황 윤 재	(인)

Abstract

Identification and Estimation of Additive Competing Risks Models with Unknown Transformation of Latent Failure Times

Oh, Narae
Department of Economics
The Graduate School
Seoul National University

This paper presents the methods of identification and estimation of additive competing risks models with unknown transformation of latent failure times. In our set up, we assume that the latent failure times are generated by nonparametric additive separable transformation regression model. The model in this paper includes a competing risk version of log-linear model, mixed proportional hazard model, accelerated failure times model, and linear transformation model. Identification of unknown additive function is accomplished using 'marginal integration' method. Our identification strategy does not depend on identification near zero, and it does not require exclusion restriction. Given our identification results, we developed uniform consistent sample analogue estimator.

keywords : Competing risks Model, Transformation Model, Identification, Marginal Integration, Estimation
Student Number : 2011-20178

Contents

1. Introduction 1

2. Identification 4

2.1 Identification of $M_k(x^\alpha)$ ·································· 4
2.2 Identification of $H_k(t)$ ···································· 10
2.3 Identification of the joint distribution ················ 11

3. Sample Analogue Estimator 13

4. Asymptotic property of Estimator 17

5. Conclusion 20

References 21

Appendix 24

1 Introduction

We consider the identification and estimation of additive competing risks models with unknown transformation of latent failure times. Suppose there are K competing risks indexed by the integer 1 to K with corresponding failure time $(T_1, \cdots, T_K)$. One observe the duration to the first failure and causes of failure, $Y = \min_k T_k$ and $\Delta = \arg\min_k T_k$, along with a covariate vector X. In Economics literature, Han and Hausman (1990) and Van den Berg et al. (2008) apply competing risks models for unemployed duration analysis, where Y is the unemployed spell and Δ indicates the reason for leaving unemployment; e.g. getting a job or giving up the job searching effort.

It is well known that the joint distribution of latent failure times is not nonparametrically identified from the joint distribution of the observed minima (Y, Δ) (Cox, 1962; Tsiatis, 1975). Heckman and Honoré (1989) and Abbring and Van den berg (2003) breaks the non-indentifiablity theorem by considering certain class of competing risks models and by assuming additional restriction that is independence of latent failure times with covariates.

The main purpose of this paper is to provide weak condition that are necessary for identification of important characteristics of competing risks model. We assume that unknown transformation of latent failure time $H_k(T_k)$ is generated by additively separable regression model.

$$\begin{aligned} H_k(T_k) &= M_k(X) + U_k \\ &\equiv \sum_{\alpha=1}^{d} M_k^{\alpha}(X^{\alpha}) + U_k, \quad k = 1, \cdots, K \end{aligned} \tag{1}$$

where H_k is unknown, differentiable, and strictly increasing function with derivative h_k, X is d-dimensional vector of covariates, X^α is α th component of X, M_k^α is unknown functions with scalar arguments with derivative $\partial M_k^\alpha(X^\alpha)/\partial X^\alpha \equiv m_k^\alpha$, and U_k is an unobserved random variable that is independent of X. It is also assumed that the distribution of U_k is unknown and the U_k may depend on each other.

The model (1) includes a large number of models that are widely used in duration analysis as special cases. These models include competing risks version of accelerated failure time model, proportional hazards model, mixed proportional hazard model. For example, the accelerated failure time model can be specified using model (1) when we assume that transformed latent failure time $H_k(T_k) = \log T_k$ (Heckman and Honoré, 1989). The mixed proportional hazard(MPH) model can be expressed with $U_k = \alpha_k + \epsilon_k$, where α_k is a cause-specific frailty term, ϵ_k is an unobserved random variable that has cumulative distribution function $F_k(\epsilon) = 1 - \exp(-e^\epsilon)$. The integrated base line hazard function is $\exp[H_k(t)]$. Abbring and Van den berg (2003) have investigated nonparametric identifiability of MPH competing risk model both in single and multiple-spell setting.

The linear transformation model has been widely used in survival analysis with a single risk, including Box-Cox (1964) regression model, proportional hazards model, and the accelerated failure time model. Horowitz (1996) developed $\sqrt{n}$-consistent and asymptotically normal estimator of unknown transformation function $H(T)$ and unknown cumulative distribution function $F(U)$. Abrevaya (1999) developed estimator of measuring effect of covariates under the fixed-effect panel version of the linear transformation model. Horowitz (1999) nonparametrically estimates baseline hazard function and distribution of unobserved heterogeneity in proportional hazard model[1]. Horowitz and Lee (2004) shows how to estimate panel data proportional hazard model with fixed effect without assuming that both base line and integrated base line hazard function belongs to known parametric family.

The model (1) is also closely related to nonparametric additive model[2]. That is, we assume that unrestricted functional form for each covariate effect is separable. Linton and Nielsen (1994) introduced "marginal integration" method to estimate the component of additive regression. Adapting additive structure for regression function mitigates curse of dimensionality, thus each of the additive component can be estimated with the one-dimensional nonparametric rate of convergence. Linton and Härdle (1996) and Horowitz (2001) extend the methods of Linton and Nielsen (1994) to a generalized additive model with a known link function and unknown link function, respectively.

Our identification strategy is closely related to that of Lee (2006) in the sense that we impose a transformation model for each latent failure times and thus consider general class of competing risks model, including proportional hazard model, MPH model, and accelerated failure times model. Lee (2006) showed that vector of parameter that measure the effect of the covariates in linear regression model can be identified up to location and scale normalization. In contrast, we generalized the identification results by allowing nonlinearlity in each covariate x^{α} for $\alpha = 1, \cdots, d$. We introduced "marginal integration" for identification and estimation of additive component of model (1). Our identification result also does not depend on identification near (or in the neighborhood of) zero, identification

[1] For more examples and application of transformation model, see Horowitz (2009, Ch.6)

[2] For more examples and application of nonparametric additive model, see Horowitz (2009, Ch.3)

at infinity, or exclusion restriction.

We also developed nonparametric estimator of $M_k^{\alpha}(x^{\alpha})$ based on our identification results. That is, we show how to estimate M_k^{α} without assuming this function belongs to a known parametric family. It will be shown that the estimator $M_{nk}^{\alpha}(x^{\alpha})$ converges almost surely as $n \to \infty$ to $M_k^{\alpha}(x^{\alpha})$ uniformly over the compact interval $[x_0, x^{\alpha}] \in S_X$.

This paper is organized as follows. In section 2, we presents the conditions that $M_k^{\alpha}(x^{\alpha})$ and $H_k(T_k)$ for $k = 1, \cdots, K$ are identified. In section 3, we propose the sample analogue kernel estimator of $M_k^{\alpha}(x^{\alpha})$ based on the identification results in section 2. The asymptotic properties of the estimator are presented in Section 4. Section 5 concludes suggesting future research.

2 Identification

2.1 Identification M_k^{α}

To begin with, observe that Equation (1) is unchanged if each component of M_k^{α} is replaced by $M_k^{\alpha} + c_{\alpha}$ for some finite constant c_{α} and H_k is replace by $\tilde{H}_k = H_k - \sum_{\alpha=1}^{d} c_{\alpha}$. Equation (1) also holds if H_k, M_k^{α}, and U_k are replace by cH_k, cM_k^{α}, and cU_k for any positive constant c. Therefore, location and scale normalizations are required to make identification possible. In this section, the location normalization is realized by setting $M_k^{\alpha}(x_0^{\alpha}) = 0$ for some constant x_0^{α} for each $k = 1, \cdots, K$. With this assumption, there exist no constant term in covariate vector X. We also assume that $d \geq 2$. Otherwise, there is nothing to left to identify since the scale factor has to be normalized.

Survival function and sub-survival function follow as

$$S(t|x) = \Pr(Y > t|X = x) \qquad Q_k(t|x) = \Pr(T_k > t, \Delta = k|X = x)$$

Also, let $p(x)$ denote the probability density function of x and S_X be the compact support of X. For $\alpha = 1, \cdots, d$ and $k = 1, \cdots, K$, define

$$A_\alpha(t, x) = \frac{\partial S(t|x)}{\partial x^\alpha} p(x) \qquad B_k(t, x) = \frac{-\partial Q_k(t|x)}{\partial t} p(x)$$

where $B(t, x) = (B_1(t, x), \cdots, B_K(t, x))'$. Note that $A_\alpha(t, x)$ and $B_k(t, x)$ is identified directly from data. In this paper, we will show that M_k^α can be expressed as functional of $A_\alpha(t, x)$ and $B_k(t, x)$ under the following assumption.[3]

Assumption 1 *(1a) U_k is independent of X with unknown probability distribution and U_k may depend on each other.*

(1b) X is $d(\geq 2)$-dimensional continuous vector of covariates with probability density function $p(x)$ that is positive on S_X except the boundary. X^α is αth component of X, X_j is jth observation of X, and X_j^α is jth observation of αth component of X.

(1c) H_k is unknown, strictly increasing, and differentiable function with derivative h_k. For $k = 1, \cdots, K$, assume that $h_k(t) \neq 0 \quad , \forall t \in S_T$.

(1d) M_k^α is unknown differentiable function with scalar argument X^α with $M_k^\alpha(x_0^\alpha) = 0$ for some $x_0^\alpha \in S_X$.

(1e) For each $k = 1, \cdots, K$

$$\int_{S_T} \frac{w_T(t)}{h_k(t)} dt = 1$$

[3]The notations and assumptions we used in this section comes from Lee (2006).

where $w_T(t)$ is scalar-value weight function on compact support S_T.

(1f) $E[B(t,X)B(t,X)'|X^\alpha = x^\alpha]$ is $k \times k$ non-singular matrix.

Assumption (1a) allows the arbitrary dependence between unobserved heterogeneity. Assumption (1b) insures the existence of probability density function of x. Location normalization can be achieved if there is no constant term in covariate vector X and if $M_k^\alpha(x_0^\alpha) = 0$ for some $x_0^\alpha \in S_X$. Assumption (1c) and (1d) specify the transformation model (Horowitz, 1996) and nonparametric additive model structure (Linton and Nielsen, 1995; Linton and Härdle, 1996), respectively. Scale normalization is accomplished by assumption (1e). This assumption is useful to create averaging effects, so the same type of scale normalization is used in transformation models (Horowitz, 1996; Lee, 2006), additive models (Linton and Härdle, 1996; Horowitz, 2001) for similar reasons. Assumption (1f) is technical assumption necessary for identification of M_k^α.

To obtain identification results of additive function M_k^α, let $M^\alpha(x^\alpha) = (M_1^\alpha(x^\alpha), \cdots, M_K^\alpha(x^\alpha)'$ for $\alpha = 1, \cdots, d$. Thus, $M^\alpha(x^\alpha)$ be the $K \times 1$ vector of unknown additive functions of α-th component of covariate vector X.

Proposition 1 *Let assumption 1 holds. Then, for each $\alpha = 1, \cdots, d$, $M^\alpha(x^\alpha)$ can be expressed as*

$$M^\alpha(x^\alpha) = \int_{x_0^\alpha}^{x^\alpha} \int_{S_T} w_T(\tau) E[B(\tau, X^\alpha, X^{(-\alpha)}) B(\tau, X^\alpha, X^{(-\alpha)})' | X^\alpha = \nu]^{-1} \times E[B(\tau, X^\alpha, X^{(-\alpha)}) A_\alpha(\tau, X^\alpha, X^{(-\alpha)}) | X^\alpha = \nu] d\tau d\nu \quad (2)$$

Proof Let $M_k(X) = \sum_{\alpha=1}^d M_k^\alpha(X^\alpha)$ and $f(u_1, \cdots, u_K)$is the joint probability density function of $(U_1, \cdots, U_K)$.

$$
\begin{aligned}
S(t|x) &= \Pr(Y > t|X = x) \\
&= \Pr(H_k(T_k) > H_k(t), \quad \forall k|X = x) \\
&= \Pr(U_k > H_k(t) - M_k(x), \quad \forall k) \\
&= \int_{H_1(t)-M_1(x)}^{\infty} \cdots \int_{H_K(t)-M_K(x)}^{\infty} f(u_1, \cdots, u_K) du_1 \cdots du_K
\end{aligned}
$$

$$
\begin{aligned}
Q_k(t|x) &= \Pr(T_k > t, \Delta = k|X = x) \\
&= \Pr(H_k(T_k) > H_k(t), H_l(T_l) > H_l(T_k), \quad \forall l \neq k|X = x) \\
&= \Pr(U_k > H_k(t) - M_k(x), U_l > H_l(T_k) - M_l(x), \quad \forall l \neq k) \\
&= \int_{H_k(t)-M_k(x)}^{\infty} \underbrace{\int_{H_1(H_k^{-1}(M_k(x)+U_k))-M_1(x)}^{\infty} \cdots \int_{H_K(H_k^{-1}(M_k(x)+U_k))-M_K(x)}^{\infty}}_{K-1 \text{ integral except k}} \\
&\qquad \times f(u_1, \cdots, u_K) \quad \underbrace{du_1 \cdots du_K}_{du_k \text{ is excluded}} \quad du_k
\end{aligned}
$$

for $k = 1, \cdots, K$. By differentiation,

$$
\begin{aligned}
\frac{\partial S(t|x)}{\partial x^\alpha} &= \sum_{k=1}^{K} m_k^\alpha(x^\alpha) \underbrace{\int_{H_1(t)-M_1(x)}^{\infty} \cdots \int_{H_K(t)-M_K(x)}^{\infty}}_{K-1 \text{ integral except k}} \\
&\qquad \times f(u_1, \cdots, u_{k-1}, H_k(t) - M_k(x), u_{k+1}, \cdots, u_K) \quad \underbrace{du_1 \cdots du_K}_{du_k \text{ is excluded}}
\end{aligned}
$$

The sub-survival function characterized by corresponding sub-density function

$$
\begin{aligned}
\frac{\partial Q_k(t|x)}{\partial t} &= -h_k(t) \underbrace{\int_{H_1(t)-M_1(x)}^{\infty} \cdots \int_{H_K(t)-M_K(x)}^{\infty}}_{K-1 \text{ integral except k}} \\
&\quad \times f(u_1, \cdots, u_{k-1}, H_k(t) - M_k(x), u_{k+1}, \cdots, u_K) \underbrace{du_1 \cdots du_K}_{du_k \text{ is excluded}}
\end{aligned}
$$

where x^α is αth component of x and $m_k^\alpha(x^\alpha)$ is derivative of $M_k(x)$ with respect to αth component of x, i.e $\partial M_k(x)/\partial x^\alpha = \partial \sum_{\alpha=1}^{d} M_k^\alpha(x^\alpha)/\partial x^\alpha \equiv m_k^\alpha(x^\alpha)$. It follows that

$$
\frac{\partial S(t|x)}{\partial x^\alpha} = \sum_{k=1}^{K} \frac{-\partial Q_k(t|x)}{\partial t} \frac{m_k^\alpha(x^\alpha)}{h_k(t)} \tag{3}
$$

Multiplying probability density function $p(x)$ to both side of equation (3) gives

$$
A_\alpha(t, x) = B(t, x)' C_\alpha(t, x^\alpha) \tag{4}
$$

where $C_\alpha(t, x^\alpha) = \left(\frac{m_1^\alpha(x^\alpha)}{h_1(t)}, \cdots, \frac{m_K^\alpha(x^\alpha)}{h_K(t)} \right)'$ and $m^\alpha(x^\alpha) = (m_1^\alpha(x^\alpha), \cdots, m_K^\alpha(x^\alpha))'$. Multiplying both side by $B(t, x)$ gives

$$
B(t, x) A_\alpha(t, x) = B(t, x) B(t, x)' C_\alpha(t, x^\alpha) \tag{5}
$$

Partition $X = (X^\alpha, X^{(-\alpha)})$, where X^α is one-dimensional direction of interest and $X^{(-\alpha)}$ is a $(d-1)$-dimensional nuisance direction, and let $x = (x^\alpha, x^{(-\alpha)})$. To identify $M_k^\alpha(x^\alpha)$, substitute $x^{(-\alpha)}$ to random variable $X^{(-\alpha)}$ and take expectation conditional on $X^\alpha = x^\alpha$. Then we can obtain

$$
E[B(t, X) A_\alpha(t, X) | X^\alpha = x^\alpha] = E[B(t, X) B(t, X)' | X^\alpha = x^\alpha] C_\alpha(t, x^\alpha) \tag{6}
$$

By the assumption that $E[B(t,X)B(t,X)'|X^\alpha = x^\alpha]$ is non-singular for every $t \in S_T$ and scale normalization,

$$m^\alpha(x^\alpha) = \int_{S_T} w(\tau) E[B(\tau,X)B(\tau,X)'|X^\alpha = x^\alpha]^{-1} \\ \times E[B(\tau,X)A_\alpha(\tau,X)|X^\alpha = x^\alpha] d\tau$$

Therefore, $M^\alpha(x^\alpha) \equiv (M_1^\alpha(x^\alpha), \cdots, M_K^\alpha(x^\alpha))'$ can be expressed as Equation (2) under the location normalization. ∎

As shown by Heckman and Honoré (1989), Abbring and Van den Berg (2003) and Lee (2006), Equation (2) also does not depend on exclusion restrictions. This implies that we can allow all the elements of covariates vector to appear in the $M_k(X)$ function, which is common feature of competing risks model.

Identification results of Heckman and Honoré (1989) and Abrring and Van den Berg (2003) are based on letting the index of time variable goes to (or near the neighborhood of)zero. Thus, the estimation methods based on the identification results only allow to use observation with failure times very close to zero (Fermanian, 2003). In contrast, the estimator proposed in Section 3 utilizes all the observed latent time variables in compact support $t \in [t_L, t_U]$. We will propose nonparametric estimator $M_n^\alpha(x^\alpha)$ based on Equation (2) in Section 3. The estimator of $M_n^\alpha(x^\alpha)$ is obtained by replacing unknown population quantities in Equation (2) with proper sample analogues.

2.2 Identification of $H_k(t)$

In this section, we present the conditions for the identification of $H_k(t)$. Note that $H_k(t)$ can be unbounded as $|t| \to \infty$, so we only consider the identification of H_k on a compact support $S_T = [t_L, t_U]$. The necessary location normalization on H_k is achieved here by assuming that there exist $t_0 \in S_T$ such that $H_k(t_0) = 0$ for each $k = 1, \cdots, K$.

Let $H(t) = (H_1, \cdots, H_K(t))'$ be vector of transformed latent failure times. To illustrate an identification result for H_k, we define element by element division operator as $a./b = (a_1/b_1, \cdots, a_K/b_K)$, where $a = (a_1, \cdots, a_K)$ and $b = (b_1, \cdots, b_K)$ be the $K-$dimensional vector, respectively.

Proposition 2 *Assumption 1 holds. Suppose $H(t_0) = 0$ for some $t_0 \in [t_L, t_U]$. Then, $H(t)$ can be expressed as*

$$H(t) = \int_{t_0}^{t} m^\alpha(x^\alpha)./\{E[B(\tau, X)B(\tau, X)'|X^\alpha = x^\alpha]^{-1} \times E[B(\tau, X)A_\alpha(\tau, X)|X^\alpha = x^\alpha]\}d\tau \quad (7)$$

for $t \in [t_L, t_U]$.

Proof From equation (6), element by element division gives

$$h(t) = m^\alpha(x^\alpha)./\{E[B(t, X)B(t, X)'|X^\alpha = x^\alpha]^{-1} \times E[B(t, X)A_\alpha(t, X)|X^\alpha = x^\alpha]\} \quad (8)$$

where $h(t) = (h_1(t), \cdots, h_K(t))'$. Thus, $H(t) = (H_1(t), \cdots, H_k(t))'$ can be expressed as equation (7) under the location normalization assumption, $H(t_0) = 0$.

The condition that all component of $m^{\alpha}(x^{\alpha})$ are nonzero is necessary. ∎

2.3 Identification of joint distribution

In this section, we consider identification of the joint distribution of $(U_1, \cdots, U_K)$ and joint survival function of $(T_1, \cdots, T_K)$ conditional on covariates. First, we define joint survival function of $(U_1, \cdots, U_K)$ as

$$S_u(u_1, \cdots, u_K) = \Pr(U_1 > u_1, \cdots, U_K > u_k)$$

In addition, we assume that when $M_k^{\alpha}(x^{\alpha})$ is identified and there is no functionally deterministic relationship among $M_K(X)$, we can exploit independent variations of one of the $M_K(X)$ given others (Lee, 2006). Define conditional survivor function given $(M_1(X), \cdots, M_K(X)) = (m_1, \cdots, m_K)$ as $S_{Y|M}(t|m_1, \cdots, m_K) = \Pr(Y > t|M_1(X) = m1, \cdots, M_K(X) = m_K)$. Note that if $M_k^{\alpha}(x^{\alpha})$ is identified for each k, we can identify $S_{Y|M}(t|m_1, \cdots, m_k)$ directly from data.

Proposition 3 *Let Assumption 1 holds. Consider $M_k(X)$ and $H(t)$ for $t \in [t_L, t_U]$ are identified. There is no functionally deterministic relationships among M_K and the support of $M = (M_1, \cdots, M_K)$ is R^K. Also, we assume that there is predetermined weight function $w_U(t)$ such that $\int w_U(t)dt = 1$ with $t \in [t_L, t_U]$. Then, for any $(u_1, \cdots, u_K)$*

$$S_U(u_1, \cdots, u_K) = \int w_U(t) S_{Y|M}(t|H_1(t) - u_1, \cdots, H_K(t) - u_K) \tag{9}$$

Proof By the definition of joint survival function of $(U_1, \cdots, U_K)$

$$S_u(u_1, \cdots, u_K) = \int_{u_1}^{\infty} \cdots \int_{u_K}^{\infty} f(u_1, \cdots, u_K) du_1 \cdots du_K$$

Also,

$$\begin{aligned} S_{Y|M}(t|m_1, \cdots, m_K) &= \Pr(U_k > H_k(t) - M_k(X) \quad \forall k) \\ &= \int_{H_1(t)-M_1(X)}^{\infty} \cdots \int_{H_K(t)-M_K(X)}^{\infty} f(u_1, \cdots, u_K) du_1 \cdots du_K \end{aligned}$$

Thus, we can express $S_U(u_1, \cdots, u_k)$ using $S_{Y|M}(t|m_1, \cdots, m_K)$ as

$$S_U(u_1, \cdots, u_k) = S_{Y|M}(t|H_1(t) - u_1, \cdots, H_K(t) - u_K)$$

for any $t \in [t_L, t_U]$. ∎

Based on the identification results of Proposition 1 to 3, we can identify distribution of latent failure times conditional on explanatory variables.

Proposition 4 *Let Assumption 1 holds. Suppose that $M_k(X)$, $H(t) fort \in [t_L, t_U]$ and $S_U(u_1, \cdots, u_K)$ for $(u_1, \cdots, u_K) \in R^K$ are identified. Then,*

$$\Pr(T_1 > t_1, \cdots, T_K > t_K|X = x) = S_U(H_1(t_1) - M_1(x), \cdots, H_K(t_K) - M_K(x)) \tag{10}$$

for any $(t_1, \cdots, t_K) \in [t_L, t_U]^K$.

3 Sample analogue Estimator

This section provides informal description of our estimator of M_k^α. The identification result proposed in Equation (2) used as a basis of the estimator proposed here. Then, the sample analogue estimator of M_k^α can be obtained replacing unknown population quantities in Equation (2) with nonparametric kernel estimator.

To begin the derivation of estimator of $M_k^\alpha(x^\alpha)$, define $S(t,x) \equiv S(t|x)p(x)$, $p_\alpha(x) \equiv \partial p(x)/\partial x^\alpha$ and $S_\alpha(t,x) \equiv \partial S(t,x)/\partial x^\alpha$. It follows that

$$\begin{aligned}
E[B(t, X^\alpha, X^{(-\alpha)})A_\alpha(t, X^\alpha, X^{(-\alpha)})|X^\alpha = x^\alpha] \\
= E[B(t, X^\alpha, X^{(-\alpha)})S_\alpha(t, X^\alpha, X^{(-\alpha)})|X^\alpha = x^\alpha] \\
- E[B(t, X^\alpha, X^{(-\alpha)})S(t|X^\alpha, X^{(-\alpha)})p_\alpha(X^\alpha, X^{(-\alpha)})|X^\alpha = x^\alpha] \\
\equiv d_1(t, x^\alpha) - d_2(t, x^\alpha)
\end{aligned}$$

Also, define $g(t, x^\alpha) \equiv E[B(t, X^\alpha, X^{(-\alpha)})B(t, X^\alpha, X^{(-\alpha)})'|X^\alpha = x^\alpha]$. Then, the Equation (2) can be rewritten as

$$M^\alpha(x^\alpha) = \int_{x_0^\alpha}^{x^\alpha} \int_{s_T} w_\tau(\tau) g(\tau, \nu)^{-1} [d_1(\tau, \nu) - d_2(\tau, \nu)] d\tau d\nu \qquad (11)$$

Equation (9) is the basis of estimator of $M^\alpha(x^\alpha)$ developed here. The estimator of $M^\alpha(x^\alpha)$ can be obtained by replacing unknown functional $g(t, x^\alpha)$, $d_1(t, x^\alpha)$, and $d_2(t, x^\alpha)$ in Equation (9) with their sample analogs $\hat{g}_n(t, x^\alpha)$, $\hat{d}_{n1}(t, x^\alpha)$, and

$\hat{d}_{n2}(t, x^\alpha)$. The resulting estimator of $M^\alpha(x^\alpha)$ is

$$M_n^\alpha(x^\alpha) = \int_{x_0^\alpha}^{x^\alpha} \int_{s_T} w_\tau(\tau) \hat{g}_n(\tau, \nu)^{-1} [\hat{d}_{n1}(\tau, \nu) - \hat{d}_{n2}(\tau, \nu)] d\tau d\nu \qquad (12)$$

Section 4 gives conditions for uniform consistency of estimator $M_n^\alpha(x^\alpha)$ over a suitable interval. Integration over (t, x^α) in Equation (12) creates averaging effect that reduce the curse of dimensionality. A similar averaging effect occurs in semiparametric estimation of transformation models (Horowitz, 1996; Lee, 2006) and additive model (Linton and Härdel, 1996; Horowitz, 2001).

The functional $g(t, x^\alpha)$, $d_1(t, x^\alpha)$, and $d_2(t, x^\alpha)$ are estimated with kernel density functions. Assume that $\{(X_i, Y_i, \Delta_i)\}_{i=1}^n$ is the random sample of (X, Y, Δ). Let X_i^α denote the αth component of X_i and $X_i^{(-\alpha)}$ be the $(d-1)$ vector of components of X_i except X_i. Let K_Y and K_1 be kernel functions of scalar argument, and K_2 be $(d-1)$ -dimensional kernel function in the sense of nonparametric estimation and regression. Let h_{ny}, h_{n1}, and h_{n2} $(n = 1, 2, \cdots)$ be the sequences of bandwidth that converges to zero as $n \to \infty$. Also, let $\mathbf{1}(\cdot)$ be the usual indicator function. Conditions that kernel functions and bandwidth parameters need to satisfy are given in section 4.

Estimate $p(x^\alpha)$, the probability density function of x^α, by

$$p_n(x^\alpha) = (nh_{n1})^{-1} \sum_{j=1}^n K_1 \left(\frac{x^\alpha - X_j^\alpha}{h_{n1}} \right) \qquad (13)$$

For each $(x^\alpha, x^{(-\alpha)}) \in S_X$, the estimator of k-th component $B_k(t, X^\alpha, X^{(-\alpha)})$

of $B(t, X^\alpha, X^{(-\alpha)})$ is obtained by

$$B_{nk}(t, x^\alpha, x^{(-\alpha)}) = (nh_{ny}h_{n1}h_{n2}^{d-1})^{-1}\sum_{j=1}^{n}\mathbf{1}(\Delta_j = k)K_Y\left(\frac{t-Y_j}{h_{ny}}\right)$$
$$\times K_1\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right)K_2\left(\frac{x^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right) \quad (14)$$

Therefore, the estimator of $g(t, x^\alpha)$ is can be expressed using Nadraya-Watson kernel estimator

$$\hat{g}_n(t, x^\alpha) = [nh_1p_n(x^\alpha)]^{-1}\sum_{i=1}^{n}B_n(t, x^\alpha, X_i^{(-\alpha)})B_n(t, x^\alpha, X_i^{(-\alpha)})'$$
$$\times K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \quad (15)$$

where $B_n(t, x^\alpha, X_i^{(-\alpha)}) = \left(B_{n1}(t, x^\alpha, X_i^{(-\alpha)}), \cdots, B_{nK}(t, x^\alpha, X_i^{(-\alpha)})\right)'$.
For each $(x^\alpha, x^{(-\alpha)}) \in S_X$, the estimator of $S_\alpha(t, X^\alpha, X^{(-\alpha)})$ is

$$S_{n\alpha}(t, x^\alpha, x^{(-\alpha)})$$
$$= (nh_{n1}^2h_{n2}^{d-1})^{-1}\sum_{j=1}^{n}\mathbf{1}(Y_j > t)K_1'\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right)K_2\left(\frac{x^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right) \quad (16)$$

The derivative of density function $p(x^\alpha, x^{(-\alpha)})$ with respect to αth component of X can be estimated by

$$p_{n\alpha}(x^\alpha, x^{-(\alpha)}) = (nh_{n1}^2h_{n2}^{d-1})^{-1}\sum_{j=1}^{n}K_1'\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right)K_2\left(\frac{x^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right) \quad (17)$$

where $K_1'(u)$ is partial derivative of kernel function with respect to α-th component of X, $K_1'(u) = \frac{\partial}{\partial X^\alpha}K(u)$, when we assume that $K_X'(u)$ exist and non-zero.

Let $p(x^\alpha, x^{(-\alpha)})$ be the joint probability density function of $(x^\alpha, x^{(-\alpha)})$. Estimate $p(x^\alpha, x^{(-\alpha)})$ by

$$p_n(x^\alpha, x^{(-\alpha)}) = (nh_{n1}h_{n2}^{d-1})^{-1} \sum_{j=1}^{n} K_1\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right) K_2\left(\frac{x^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right) \quad (18)$$

The estimator of survival function $S(t|X^\alpha, X^{(-\alpha)})$ is

$$\begin{aligned} S_n(t|x^\alpha, x^{(-\alpha)}) = [nh_{n1}h_{n2}^{d-1}p_n(x^\alpha, x^{(-\alpha)})]^{-1} \sum_{j=1}^{n} \mathbf{1}(Y_j > t) \\ \times K_1\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right) K_2\left(\frac{x^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right) \quad (19) \end{aligned}$$

The estimator of $d_1(t, x^\alpha)$ and $d_2(t, x^\alpha)$ are obtained respectively by

$$\begin{aligned} \hat{d}_{n1}(t, x^\alpha) = [nh_{n1}p_n(x^\alpha)]^{-1} \sum_{i=1}^{n} B_n(t, x^\alpha, X_i^{(-\alpha)}) S_{n\alpha}(t, x^\alpha, X_i^{(-\alpha)}) \\ \times K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \quad (20) \end{aligned}$$

$$\begin{aligned} \hat{d}_{n2}(t, x^\alpha) = [nh_{n1}p_n(x^\alpha)]^{-1} \sum_{i=1}^{n} B_n(t, x^\alpha, X_i^{(-\alpha)}) S_n(t|x^\alpha, X_i^{(-\alpha)}) \\ \times p_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \quad (21) \end{aligned}$$

Therefore, the estimator $M_n^\alpha(x^\alpha)$ of $M^\alpha(x^\alpha)$ is obtained by substituting Equation (15), (20) and (21) into (12).

4 Asymptotic properties of the estimator

This section provides the assumptions used in proving uniform consistency of $M_n^\alpha(x^\alpha)$.[4]

Assumption 2 *(Sampling)* $\{(X_i, Y_i, \Delta_i)\}_{i=1}^n$ *is a random sample of* (X, Y, Δ).

Let $p(t, x^\alpha, x^{(-\alpha)})$ be the joint probability density function of $(t, x^\alpha, x^{(-\alpha)})$.

Assumption 3 *(Smoothness) The distribution of* $(t, x^\alpha, x^{(-\alpha)})$ *is absolutely continuous with respect to Lebesque measure. In addition, there are open intervals of the real line,* I_T *and* I_X*, such that*
(a) $I_T = [0, \tau_T]$, *where* $\tau_T < \infty$ *and* I_X *is open.*
(b) $p(t, x^\alpha, x^{(-\alpha)})$ *and* $p(t|x^\alpha, x^{(-\alpha)})$ *are bounded uniformly over* $(t, x^\alpha, x^{(-\alpha)}) \in I_T \times I_X$*. Moreover,* $\inf\{p(t, x^\alpha, x^{(-\alpha)}), (t, x^\alpha, x^{(-\alpha)}) \in I_T \times I_X\} > 0$.
(c) $B_k(t, x^\alpha, x^{(-\alpha)})$, $S(t, x^\alpha, x^{(-\alpha)})$, $S(t|x^\alpha, x^{(-\alpha)})$, $p(t, x^\alpha, x^{(-\alpha)})$, $S_\alpha(t, x^\alpha, x^{(-\alpha)})$, *and* $p_\alpha(t, x^\alpha, x^{(-\alpha)})$ *are q-times continuously differentiable with respect to* $t \in S_T$*, and r-times continuously differentiable with respect to* $x^\alpha \in S_X$.

This assumption insures that $g(t, x^\alpha)$, $d_1(t, x^\alpha)$ and $d_2(t, x^\alpha)$ exist and that the denominator in Equation (11) is bounded away from zero. The existence of higher order derivative of joint density function $p(\cdot)$ is needed to insure that the bias terms associated with kernel estimation of $\hat{g}_n$, $\hat{d}_{n1}$ and $\hat{d}_{n2}$ vanish sufficiently fast.

[4]The assumption used for proving rate of convergence and asymptotic normality of M_n^α is stronger than the assumption we suggested in this section. Especially, to make the certain bias and remainder terms goes to zero as $n \to \infty$, K_1 must be higher order kernel because $\hat{d}_{n1}$ and $\hat{d}_{n2}$ are functional of derivative of K_1 (See equation (16) and (17)). Derivative functional converge relatively slowly so the higher order kernel for K_1 is needed to insure sufficiently rapid convergence.

Assumption 4 *(Weight Function) (a) For each* $k = 1, \cdots, K$, $\int_{S_T} w_T(t)/h_k(t)dt = 1$. *(b) The weight function* $w_t(\cdot)$ *is bounded, non-negative function in the compact support* $S_T \subset I_T$. *(c)* $w_T(\cdot)$ *is q-times continuously differentiable for all* $t \in S_T$.

Assumption 5 *(Kernels)* K_Y *has support* $[-1, 1]$, *bounded and symmetrical about 0, has bounded variation, and satisfies*

$$\int_{-1}^{1} u^j K_Y(u)du = \begin{cases} 1 & \text{if} \quad j = 0, \\ 0 & \text{if} \quad l \le j \le q-1, \\ A_Y < \infty & \text{if} \quad j = q. \end{cases}$$

where A_Y *is positive constant.*

K_1 *has a support* $[-1, 1]$, *bounded and symmetrical about 0, has bounded variation and satisfies*

$$\int_{-1}^{1} u^j K_1(u)du = \begin{cases} 1 & \text{if} \quad j = 0, \\ 0 & \text{if} \quad l \le j \le r-1, \\ A_1 < \infty & \text{if} \quad j = r. \end{cases}$$

where A_1 *is positive constant.*

K_2 *is* $(d-1)$ *dimensional product of univariate kernel of order s,* $K_2 = \Pi_{j=1}^{d-1} K(u_j)$. K_2 *has a support* $[-1, 1]$, *bounded and symmetrical about 0, has bounded variation* *and satisfies*

$$\int_{-1}^{1} u^j K_2(u)du = \begin{cases} 1 & if \quad j = 0, \\ 0 & if \quad l \leq j \leq s-1, \\ A_2 < \infty & if \quad j = s. \end{cases}$$

where A_2 is positive constant.

K_1 is everywhere differentiable. The derivatives $K_1'(u) = dK_1(u)/du$ is bounded, Lipschitz continuous, and has bounded variation.

Müller (1984) provides kernel satisfying above assumptions.

Assumption 6 *(bandwidth) As $n \to \infty$, $h_{ny} \to 0$, $h_{n1} \to 0$, $h_{n2} \to 0$, and* $\log n/(nh_{ny}h_{n1}^3 h_{n2}^{d-1})^{1/2} \to 0$.

The following theorem gives the main results of this section.

Theorem 1 *Let Assumptions 1-6 holds. Then for each $\alpha = 1, \cdots, d$,*

$$\underset{n\to\infty}{plim} \sup_{x^\alpha} |M_n^\alpha(x^\alpha) - M_\alpha(x^\alpha)| = 0$$

Theorem 1 is proved in two step whose details are given in mathematical appendix A: First, we will use Taylor series expansion to approximate $\hat{g}_n^{-1}(\hat{d}_{n1} - \hat{d}_{n2})$ uniformly over $(t, x) \in S_T \times S_X$ by a linear functional kernel estimator. Second, combining a uniform law of large numbers of Pollard(1984) and standard arguments in kernel estimation show that the quantity observed by replacing

$\hat{g}_n^{-1}(\hat{d}_{n1} - \hat{d}_{n2})$ in Equation (12) with the linear approximation converges almost surely to $M_\alpha(x^\alpha)$ uniformly over the proper compact support.

5 Conclusion

This paper has shown the methods for identification and estimation of additive competing risks models with unknown transformation of latent failure times. The unknown function $M_k^\alpha(x^\alpha)$ identified using 'marginal integration' method. Based on this result, we also suggest the identification strategy of unknown transformation function, joint survival function of $(U_1, \cdots, U_K)$ and joint survival function of $(T_1, \cdots, T_K)$ conditional on explanatory variables. The sample analogue estimator of unknown additive function is uniformly consistent over the given compact support. Models that can be identified and estimated with new method include competing risks version of log-linear model, mixed proportional hazards model, accelerated failure times model. Presenting the asymptotic normality of the estimator applying uniform result of degenerate U-process is remained as future research.

References

[1] Abbring, J.H. and Van den Berg, G.J. (2003), "The identifiablity of the Mixed Proportional Hazards Competing Risks Model", *J.R. Statist. Soc. B*, 65, 701-710.

[2] Abrevaya, J. (1999), " Leapfrog estimation of a fixed-effect model with unknown transformation of dependent variable", *Journal of Econometrics*, 93, 203-228

[3] G. E. P. Box and D. R. Cox (1964), " An Analysis of Transformations", *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 26, No. 2 , 211-252.

[4] Clayton, D. and Cuzick, J. (1985),"Multivariate Generalizations of the Proportional Hazards Model", *J.R. Statist. Soc. A*, 148, 82-117.

[5] Fan, J. Härdle, W. and Mammen, E. (1998),"Direct Estimation of Low-Dimensional Components in Additive Models", *The Annals of Statistics*, 26, 943-971.

[6] Fermanian, J.D. (2003), "Nonparametric Estimation of Competing Risks Models with Covariates", *Journal of Multivariate Analysis*, 85, 156-191.

[7] Flinn, C.J. and Heckman, J.J. (1983),"Are Unemployment and Out of Labor Force Behaviorally Distinct Labor Force States?", *Journal of Labor Economics*, 1, 28-42.

[8] Han, A. and Hausman, J.A. (1990),"Flexible Parametric Estimation of Duration and Competing Risk Models", *Journal of Applied Econometrics*, 5, 1-28.

[9] Härdle, W. and Linton, O. (1994), *Applied Nonparametrics Methods, Handbook of Econometrics(edited by R.F. Engle and D.L. MaFadden)*, Vol.IV, Elsevier Science.

[10] Heckman, J.J. and Honoré, B.E. (1989),"The Identifiablity of the Competing Risks Model", *Biometrika*, 76, 325-330.

[11] Honoré, B.E. (1993), "Identification Results for Duration Models with Multiple Spells", *The Review of Economic Studies*, 60, 241-246.

[12] Horny, G. and Picchio, M. (2010), "Identification of lagged duration dependence in multiple-spell competing risks models", *Economics Letters*, 106, 241-243.

[13] Horowitz, L.J. (1996), "Semiparametric Estimation of Regression Model with an Unknown Transformation of the Dependent Variable", *Econometrica*, 64, 103-137.

[14] Horowitz, L.J. (1999), "Semiparametric Estimation of Proportional Hazard Model with Unobserved Heterogeneity", *Econometrica*, 67, 1001-1028.

[15] Horowitz, L.J. (2001), "Nonparametric Estimation of a Generalized Additive Model with an Unknown Link Function", *Econometrica*, 69, 499-513.

[16] Horowitz, L.J. and Lee, S. (2004), "Semiparametric Estimation of a Panel Data Proportional Hazards Model with Fixed Effects", *Journal of Econometrics*, 119, 155-198.

[17] Lee, S. (2006), "Identification of a Competing Risks Model with Unknown Transformation of Latent Failure Times", *Biometrika*, 93, 996-1002.

[18] Lee, S. and Lewbel A. (2013) “Nonparametric Identification of Accelerated Failure Time Competing Risks Models”, *Econometric Theory*, *Forthcoming*

[19] Linton, O.B and Härdle, W. (1996), “Estimation of Additive Regression Models with Known Links”, *Biometrika*, 83, 529-540.

[20] Linton, O.B. and Nielsen, J.P. (1995), “A Kernel Method of Estimation Structure Nonparametric Regression Based on Marginal Integration”, *Biometrika*, 82, 93-100.

[21] Pakes, A. and Pollard, D. (1989), “Simulation and the Asymptotics of Optimization Estimators”, *Econometrica*, 57, 1027-1057.

[22] Pollard, D. (1984), “Convergence of Stochastic Process”, *Springer*.

[23] Tsiatis, A. (1975), “A Nonidentifiablity Aspect of the Problem of Competing Risks”, *Proc. Nat. Acad. Sci.*,72, 20-22.

[24] Van den Berg, G.J., van Lomwel, A.G.C and van Ours, J.C. (2008), “Nonparametric estimation of a dependent competing risks model for unemployment durations”, *Empirical Economics*, 34, 477-491.

A Mathematical Appendix: Proof of Theorem 1

Define $G(t, x^\alpha) = g(t, x^\alpha)p(x^\alpha)$, $D_1(t, x^\alpha) = d_1(t, x^\alpha)p(x^\alpha)$, and $D_2(t, x^\alpha) = d_2(t, x^\alpha)p(x^\alpha)$. Then, $\hat{G}_n(t, x^\alpha) = \hat{g}_n(t, x^\alpha)p_n(x^\alpha)$, $\hat{D}_{n1}(t, x^\alpha) = \hat{d}_{n1}(t, x^\alpha)p_n(x^\alpha)$, and $\hat{D}_{n2}(t, x^\alpha) = \hat{d}_{n2}(t, x^\alpha)p_n(x^\alpha)$. Equation (10) holds if $\hat{g}_n$, $\hat{d}_{n1}$, and $\hat{d}_{n2}$ is replaced with $\hat{G}_n$, $\hat{D}_{n1}$, and $\hat{D}_{n2}$. It is more convenient to use $\hat{G}_n$, $\hat{D}_{n1}$, and $\hat{D}_{n2}$ than $\hat{g}_n$, $\hat{d}_{n1}$, and $\hat{d}_{n2}$ for proving Theorem 1. Euclidean class of functions is defined as in Pakes and Pollard (1989).

Lemma 1

$$\begin{aligned} B_n(t, x^\alpha, X_i^{(-\alpha)}) &= B(t, x^\alpha, X_i^{(-\alpha)}) \\ &\quad + O(h_{ny}^q) + O(h_{n1}^r) + O(h_{n2}^s) + o(\log n/((nh_{ny}h_{n1}h_{n2}^{d-1})^{1/2}) \end{aligned}$$

almost surely uniformly over $(t, x^\alpha, X_i^{(-\alpha)}) \in S_T \times S_X$.

Proof By example (2.10) and Lemma 2.14 of Pakes and Pollard (1989, p.1035), summand of $B_{nk}(t, x^\alpha, X_i^{(-\alpha)})$ is Euclidean. In addition,

$$\begin{aligned} & E\left[\mathbf{1}(\Delta_j = k)K_Y\left(\frac{t - y_j}{h_{ny}}\right)K_1\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right)K_2\left(\frac{X_i^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right)\right]^2 \\ &= \int \mathbf{1}(\Delta = k)K_Y^2\left(\frac{t - \varphi}{h_{ny}}\right)K_1^2\left(\frac{x^\alpha - \xi^\alpha}{h_{n1}}\right)K_2^2\left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}}\right) \\ &\qquad \times p(\varphi|\xi^\alpha, \xi^{(-\alpha)})p(\xi^\alpha, \xi^{(-\alpha)})d\varphi d\xi^\alpha d\xi^{(-\alpha)} \\ &= h_{ny}h_{n1}h_{n2}^{(d-1)}\int \mathbf{1}(\Delta = k)K_Y^2(u)K_1^2(v)K_2^2(w) \\ &\qquad \times p(t - h_{ny}u|x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) \\ &\qquad \times p(x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w)dudvdw \\ &\leq M_B h_{ny}h_{n1}h_{n2}^{(d-1)} \end{aligned}$$

for some finite constant M_b. Applying Theorem 2.37 of Pollard (1984) yields

$$\sup_{t, x^\alpha, X_i^{(-\alpha)}} |B_{nk}(t, x^\alpha, X_i^{(-\alpha)}) - EB_{nk}(t, x^\alpha, X_i^{(-\alpha)})| = o(\log n/(nh_{ny}h_{n1}h_{n2}^{(d-1)})^{1/2}) \quad (22)$$

almost surely as $n \to \infty$.

In addition,

$$
\begin{aligned}
&EB_{nk}(t, x^\alpha, X_i^{(-\alpha)}) \\
&= (nh_{ny}h_{n1}h_{n2}^{(d-1)})^{-1} \int \mathbf{1}(\Delta = k) K_Y\left(\frac{t-\varphi}{h_{ny}}\right) K_1\left(\frac{x^\alpha - \xi^\alpha}{h_{n1}}\right) \\
&\qquad \times K_2\left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}}\right) p(\varphi|\xi^\alpha, \xi^{(-\alpha)}) p(\xi^\alpha, \xi^{(-\alpha)}) d\varphi d\xi^\alpha d\xi^{(-\alpha)} \\
&= -\int \mathbf{1}(\Delta = k) K_Y(u) K_1(v) K_2(w) p(t - h_{ny}u | x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) \\
&\qquad\qquad \times p(x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) dudvdw \\
&= B_k(t, x^\alpha, X_i^{(-\alpha)}) + O(h_{ny}^q) + O(h_{n1}^r) + O(h_{n2}^s) \qquad (23)
\end{aligned}
$$

uniformly over $(t, x^\alpha, X_i^{(-\alpha)}) \in S_T \times S_X$. ∎

Lemma 2 *Under the assumption 2-5,*

$$
S_{n\alpha}(t, x^\alpha, X_i^{(-\alpha)}) = S_\alpha(t, x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s) + o(\log n/(nh_{n1}^3 h_{n2}^{(d-1)})^{1/2})
$$

almost surely uniformly over $(t, x^\alpha, X_i^{(-\alpha)}) \in S_T \times S_X$.

Proof By example (2.10) and Lemma 2.14 of Pakes and Pollard (1989), the summand of $S_{n\alpha}(t, x^\alpha, X_i^{(-\alpha)})$ is Euclidean.

$$
\begin{aligned}
&E\left[\mathbf{1}(Y_j > t) K_1'\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right) K_2\left(\frac{X_i^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right)\right]^2 \\
&= \int \mathbf{1}(\varphi > t) K_1'^2\left(\frac{x^\alpha - \xi^\alpha}{h_{n1}}\right) K_2^2\left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}}\right) \\
&\qquad\qquad \times p(\varphi|\xi^\alpha, \xi^{(-\alpha)}) p(\xi^\alpha, \xi^{(-\alpha)}) d\varphi d\xi^\alpha \xi^{(-\alpha)} \\
&= h_{n1}h_{n2}^{(d-1)} \int \mathbf{1}(u > t) K_1'^2(v) K_2^2(w) p(u | x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) \\
&\qquad\qquad \times p(x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) dudvdw \\
&\le M_{S_\alpha} h_{n1} h_{n2}^{(d-1)}
\end{aligned}
$$

for some finite constant M_{S_α}. By Theorem 2.37 of Pollard (1984),

$$
\sup_{t, x^\alpha, X_i^{(-\alpha)}} |S_{n\alpha}(t, x^\alpha, X_i^{(-\alpha)}) - ES_{n\alpha}(t, x^\alpha, X_i^{(-\alpha)})| = o(\log n/(nh_{n1}^3 h_{n2}^{(d-1)})^{1/2}) \qquad (24)
$$

almost surely as $n \to \infty$. In addition,

$$
\begin{aligned}
&ES_{n\alpha}(t, x^\alpha, X_i^{(-\alpha)}) \\
&= (h_{n1}^2 h_{n2}^{(d-1)})^{-1} \int \mathbf{1}(\varphi > t) K_1' \left(\frac{x^\alpha - \xi^\alpha}{h_{n1}} \right) K_2 \left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}} \right) \\
&\qquad \times p(\varphi | \xi^\alpha, \xi^{(-\alpha)}) p(\xi^\alpha, \xi^{(-\alpha)}) d\varphi d\xi^\alpha \xi^{(-\alpha)} \\
&= \int \mathbf{1}(u > t) K_1'(v) K_2(w) p(u | x^\alpha - h_{n1} v, X_i^{(-\alpha)} - h_{n2} w) \\
&\qquad \times p(x^\alpha - h_{n1} v, X_i^{(-\alpha)} - h_{n2} w) du dv dw \\
&= S_\alpha(t, x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s) \qquad (25)
\end{aligned}
$$

uniformly over $(t, x^\alpha, X_i^{(-\alpha)}) \in S_T \times S_X$ ∎

Lemma 3 *Under the assumption 2-5,*

$$
p_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) = p_\alpha(x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^q) + O(h_{n2}^r) + o(\log n / (n h_{n1}^3 h_{n2}^{(d-1)})^{1/2})
$$

almost surely uniformly over $(x_\alpha, X_i^{(-\alpha)}) \in S_X$

Proof By example (2.10) and Lemma 2.14 of Pakes and Pollard (1989), the summand of $p_{n\alpha}(x^\alpha, X_i^{(-\alpha)})$ is Euclidean.

$$
\begin{aligned}
&E \left[K_1' \left(\frac{x_\alpha - X_j \alpha)}{h_{n1}} \right) K_2 \left(\frac{X_i^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}} \right) \right]^2 \\
&= \int K_1^{'2} \left(\frac{x_\alpha - \xi^\alpha}{h_{n1}} \right) K_2^2 \left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}} \right) p(\xi^\alpha, \xi^{(-\alpha)}) d\xi^\alpha d\xi^{(-\alpha)} \\
&= h_{n1} h_{n2}^{d-1} \int K_1^{'2}(v) K_2^2(w) p(x^\alpha - h_{n1} v, X_i^{(-\alpha)} - h_{n2} w) dv dw \\
&\leq M_{p_\alpha} h_{n1} h_{n2}^{(d-1)}
\end{aligned}
$$

Applying Theorem 2.37 of Pollard (1984) yields,

$$
\sup_{t, x^\alpha, X_i^{(-\alpha)}} |p_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) - E p_{n\alpha}(x^\alpha, X_i^{(-\alpha)})| = o(\log n / (n h_{n1}^3 h_{n2}^{(d-1)})^{1/2}) \qquad (26)
$$

almost surely as $n \longrightarrow \infty$. In addition,

$$\begin{aligned}
Ep_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) &= (h_{n1}h_{n2}^{(d-1)})^{-1}\int K_1'\left(\frac{x_\alpha-\xi^\alpha}{h_{n1}}\right)K_2\left(\frac{X_i^{(-\alpha)}-\xi^{(-\alpha)}}{h_{n2}}\right) \\
&\qquad \times p(\xi^\alpha,\xi^{(-\alpha)})d\xi^\alpha d\xi^{(-\alpha)} \\
&= (h_{n1})^{-1}\int K_1'(v)K_2(w)p(x^\alpha-h_{n1}v, X_i^{(-\alpha)}-h_{n2}w)dvdw \\
&= \int K_1(v)K_2(w)\frac{\partial}{\partial x^\alpha}p(x^\alpha-h_{n1}v, X_i^{(-\alpha)}-h_{n2}w)dvdw \\
&= p_\alpha(x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s) \qquad (27)
\end{aligned}$$

uniformly over $(x^\alpha, X_i^{(-\alpha)}) \in S_X$ ∎

Define

$$S_{n1}(t,x^\alpha,X_i^{(-\alpha)}) = (nh_1h_2^{(d-1)})^{-1}\sum_{j=1}^n \mathbf{1}(Y_j>t)K_1\left(\frac{x^\alpha-X_j^\alpha}{h_{n1}}\right) \times K_2\left(\frac{X_i^{(-\alpha)}-X_j^{(-\alpha)}}{h_{n2}}\right) \qquad (28)$$

Lemma 4 *Under the assumption 2-5,*

$$S_{n1}(t,x^\alpha,X_i^{(-\alpha)}) = S_1(t,x^\alpha,X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s) + o(\log n/(nh_{n1}h_{n2}^{(d-1)})^{1/2})$$

almost surely uniformly over $(t,x^\alpha,X_i^{(-\alpha)}) \in S_T \times S_X$.

Proof By example (2.10) and Lemma 2.14 of Pakes and Pollard (1989, p.1035), the summand of $S_n(t,x^\alpha,X_i^{(-\alpha)})$ is Euclidean.

$$\begin{aligned}
&E\left[\mathbf{1}(Y_j>t)K_1\left(\frac{x^\alpha-X_j^\alpha}{h_{n1}}\right)K_2\left(\frac{X_i^{(-\alpha)}-X_j^{(-\alpha)}}{h_{n2}}\right)\right]^2 \\
&= \int \mathbf{1}(\varphi>t)K_1^2\left(\frac{x^\alpha-\xi^\alpha}{h_{n1}}\right)K_2^2\left(\frac{X_i^{(-\alpha)}-\xi^{(-\alpha)}}{h_{n2}}\right)p(\varphi,\xi^\alpha,\xi^{(-\alpha)})d\varphi d\xi^\alpha d\xi^{(-\alpha)} \\
&= h_{n1}h_{n2}^{(d-1)}\int \mathbf{1}(u>t)K_1^2(v)K_2^2(w)p(u,x^\alpha-h_{n1}v,X_i^{(-\alpha)}-h_{n2}w)dudvdw \\
&\leq M_{S_1}h_{n1}h_{n2}^{(d-1)}
\end{aligned}$$

for some finite M_{S_1}. Applying Theorem 2.37 of Pollard (1984) yields,

$$\sup_{t,x^\alpha,X_i^{(-\alpha)}} |S_{n1}(t,x^\alpha,X_i^{(-\alpha)}) - ES_{n1}(t,x^\alpha,X_i^{(-\alpha)})| = o(\log n/(nh_{n1}h_{n2}^{(d-1)})^{1/2}) \qquad (29)$$

In addition,

$$
\begin{aligned}
&ES_{n1}(t, x^\alpha, X_i^{(-\alpha)}) \\
&= (h_{n1}h_{n2}^{d-1})^{-1} \int \mathbf{1}(\varphi > t) K_1\left(\frac{x^\alpha - \xi^\alpha}{h_{n1}}\right) K_2\left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times p(\varphi, \xi^\alpha, \xi^{(-\alpha)}) d\varphi d\xi^\alpha d\xi^{(-\alpha)} \\
&= \int \mathbf{1}(y > t) K_1(v) K_2(w) p(u, x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) du dv dw \\
&= S_1(t, x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s)
\end{aligned} \tag{30}
$$

uniformly over $(t, x^\alpha, X_i^{(-\alpha)}) \in S_T \times S_X$. ▮

Lemma 5 *Under the assumption 2-5,*

$$
P_n(x^\alpha, X_i^{(-\alpha)}) = P(x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s) + o(\log n/(nh_{n1}h_{n2}^{(d-1)})^{1/2})
$$

almost surely uniformly over $(x^\alpha, X_i^{(-\alpha)}) \in S_X$

Proof By example (2.10) and Lemma 2.14 of Pakes and Pollard (1989, p.1035), the summand of $P_n(x^\alpha, X_i^{(-\alpha)})$ is Euclidean.

$$
\begin{aligned}
&E\left[K_1\left(\frac{x^\alpha - X_j^\alpha}{h_{n1}}\right) K_2\left(\frac{X_i^{(-\alpha)} - X_j^{(-\alpha)}}{h_{n2}}\right)\right]^2 \\
&= \int K_1^2\left(\frac{x^\alpha - \xi^\alpha}{h_{n1}}\right) K_2^2\left(\frac{X_i^{(-\alpha)} - \xi^{(-\alpha)}}{h_{n2}}\right) p(\xi^\alpha, \xi^{(-\alpha)}) d\xi^\alpha d\xi^{(-\alpha)} \\
&= \int K_1^2(v) K_2^2(w) p(x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w) dv dw \\
&\leq M_P h_{n1} h_{n2}^{(d-1)}
\end{aligned}
$$

for some finite constant M_P. Applying Theorem 2.37 of Pollard (1984) yields,

$$
\sup_{x^\alpha, X_i^{(-\alpha)}} |P_n(x^\alpha, X_i^{(-\alpha)}) - EP_n(x^\alpha, X_i^{(-\alpha)})| = o(\log n/(nh_{n1}h_{n2}^{(d-1)})^{1/2}) \tag{31}
$$

almost surely as $n \to \infty$. In addition,

$$\begin{aligned}
&EP_n(x^\alpha, X_i^{(-\alpha)}) \\
&= (h_{n1}h_{n2}^{(d-1)})^{-1}\int K_1\left(\frac{x^\alpha-\xi^\alpha}{h_{n1}}\right)K_2\left(\frac{X_i^{(-\alpha)}-\xi^{(-\alpha)}}{h_{n2}}\right)p(\xi^\alpha,\xi^{(-\alpha)})d\xi^\alpha d\xi^{(-\alpha)} \\
&= \int K_1(v)K_2(w)P(x^\alpha - h_{n1}v, X_i^{(-\alpha)} - h_{n2}w)dvdw \\
&= P(x^\alpha, X_i^{(-\alpha)}) + O(h_{n1}^r) + O(h_{n2}^s) \qquad (32)
\end{aligned}$$

uniformly over $(x^\alpha, X_i^{(-\alpha)}) \in S_X$. ∎

Define

$$\begin{aligned}
\hat{D}_{n1}(t,x^\alpha) &= \frac{1}{nh_{n1}}\sum_{i=1}^n B_n(t,x^\alpha,X_i^{(-\alpha)})S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha-X_i^\alpha}{h_{n1}}\right) \\
D_{n1}(t,x^\alpha) &= \frac{1}{nh_{n1}}\sum_{i=1}^n B(t,x^\alpha,X_i^{(-\alpha)})S_\alpha(t,x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha-X_i^\alpha}{h_{n1}}\right) \\
D_1(t,x^\alpha) &= \int B(t,x^\alpha,X^{(-\alpha)})S_\alpha(t,x^\alpha,X^{(-\alpha)})P(t,x^\alpha,X^{(-\alpha)})dtdX^{(-\alpha)} \\
P_n^*(x^\alpha) &= \frac{1}{nh_{n1}}\sum_{i=1}^n |K_1\left(\frac{x^\alpha-X_i^\alpha}{h_{n1}}\right)|
\end{aligned}$$

The k-th components of $\hat{D}_{1n}(t,x^\alpha)$ and $D_{1n}(t,x^\alpha)$ can be written as

$$\begin{aligned}
\hat{D}_{n1k}(t,x^\alpha) &= \frac{1}{nh_1}\sum_{i=1}^n B_{nk}(t,x^\alpha,X_i^{(-\alpha)})S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha-X_i^\alpha}{h_{n1}}\right) \\
D_{n1k}(t,x^\alpha) &= \frac{1}{nh_1}\sum_{i=1}^n B_k(t,x^\alpha,X_i^{(-\alpha)})S_\alpha(t,x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha-X_i^\alpha}{h_{n1}}\right) \\
D_1(t,x^\alpha) &= \int B_k(t,x^\alpha,X^{(-\alpha)})S_\alpha(t,x^\alpha,X^{(-\alpha)})P(t,x^\alpha,X^{(-\alpha)})dtdX^{(-\alpha)}
\end{aligned}$$

Lemma 6 *Under the assumption 1-4,*

$$\hat{D}_{n1}(t,x^\alpha) = D_1(t,x^\alpha) + O(h_{n1}^r) + o(\log n/(nh_{n1})^{1/2})$$

almost surely uniformly over $(t,x^\alpha) \in S_t \times S_X$

Proof Since almost surely convergence implies the convergence in probability, the uniform rate also holds in probability. Combining the results in Lemma 1 to Lemma 5 with the conditions that $nh_{n1}h_{n2}^{d-1}/\log n \to \infty$ and $nh_1^3 h_2^{d-1}/\log n \to \infty$ as $n \to \infty$, we obtain

$$\sup_{t,x^\alpha,X_i^{(-\alpha)}} |B_{nk}(t,x^\alpha,X_i^{(-\alpha)}) - B_k(t,x^\alpha,X_i^{(-\alpha)})| = o_p(1)$$
$$\sup_{t,x^\alpha,X_i^{(-\alpha)}} |S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)}) - S_\alpha(t,x^\alpha,X_i^{(-\alpha)})| = o_p(1)$$
$$\sup_{t,x^\alpha,X_i^{(-\alpha)}} |p_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)}) - p_\alpha(t,x^\alpha,X_i^{(-\alpha)})| = o_p(1)$$
$$\sup_{t,x^\alpha,X_i^{(-\alpha)}} |S_{n1}(t,x^\alpha,X_i^{(-\alpha)}) - S_1(t,x^\alpha,X_i^{(-\alpha)})| = o_p(1)$$
$$\sup_{x^\alpha,X_i^{(-\alpha)}} |p_n(x^\alpha,X_i^{(-\alpha)}) - p(x^\alpha,X_i^{(-\alpha)})| = o_p(1)$$

$B_k(t,x^\alpha,X_i^{(-\alpha)})$ and $S_\alpha(t,x^\alpha,X_i^{(-\alpha)})$ are bounded in compact support $S_T \times S_X$ and linear approximation of $B_{nk}S_{n\alpha} - B_k S_\alpha$ gives

$$\begin{aligned}
&|B_{nk}(t,x^\alpha,X_i^{(-\alpha)})S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)}) - B_k(t,x^\alpha,X_i^{(-\alpha)})S_\alpha(t,x^\alpha,X_i^{(-\alpha)})|\\
&\le |S_\alpha(t,x^\alpha,X_i^{(-\alpha)})||B_{nk}(t,x^\alpha,X_i^{(-\alpha)}) - B_k(t,x^\alpha,X_i^{(-\alpha)})|\\
&+|S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)}) - S_\alpha(t,x^\alpha,X_i^{(-\alpha)})||S_\alpha(t,x^\alpha,X_i^{(-\alpha)})|\\
&+R_{n1}(t,x^\alpha,X_i^{(-\alpha)})
\end{aligned}$$

so that $|B_{nk}(t,x^\alpha,X_i^{(-\alpha)})S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)}) - B_k(t,x^\alpha,X_i^{(-\alpha)})S_\alpha(t,x^\alpha,X_i^{(-\alpha)})| = o_p(1)$ uniformly over $(t,x^\alpha,X_i^{(-\alpha)}) \in S_T \times S_X$, where the remainder term $R_{n1} = O[|B_{nk} - B_k||S_{n\alpha} - S_\alpha|]$ goes to zero as $n \to \infty$.

We can easily show that summand of P_n^* is Euclidean by example (2.10) and Lemma (2.14) of Pakes and Pollard (1989). Thus, Applying Theorem 2.37 of Pollard (1984) yields,

$$\sup_{x^\alpha}|P_n^*(x^\alpha) - EP_n^*(x^\alpha)| = o(\log n/(nh_{n1}^{1/2})) \tag{33}$$

In addition,

$$\begin{aligned}
&EP_n^*(x^\alpha)\\
=\ &\frac{1}{h_{n1}}\int |K_1\left(\frac{x^\alpha-\xi^\alpha}{h_{n1}}\right)p(\xi^\alpha)d\xi^\alpha\\
=\ &\int |K_1(v)|p(x^\alpha - h_{n1}v)dv\\
=\ &O(1)
\end{aligned} \tag{34}$$

uniformly over $x^\alpha \in S_X$. Thus,

$$
\begin{aligned}
&|\hat{D}_{n1k}(t,x^\alpha) - D_{n1k}(t,x^\alpha)| \\
&\leq \sup|B_{nk}(t,x^\alpha,X_i^{(-\alpha)})S_{n\alpha}(t,x^\alpha,X_i^{(-\alpha)}) - B_k(t,x^\alpha,X_i^{(-\alpha)})S_\alpha(t,x^\alpha,X_i^{(-\alpha)})| \\
&\times(nh_{n1})^{-1}\sum_{i=1}^{n}|K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right)| \\
&= o_p(1)O(1) \\
&= o_p(1) \qquad (35)
\end{aligned}
$$

uniformly over $(t,x^\alpha) \in S_T \times S_X$.

(Part B) As in Lemma 1, $K_1\left(\frac{x^\alpha-\xi^\alpha}{h_{n1}}\right)$ is Euclidean, and $B_k(t,x^\alpha,X^{(-\alpha)})$ and $S_\alpha(t,x^\alpha,X^{(-\alpha)})$ are Euclidean because these are single functions. Hence, by Lemma 2.14 of Pakes and Pollard (1989), the summand of D_{n1} is Euclidean. Applying Theorem 2.37 of Pollard (1984) yields,

$$
\sup_{t,x^\alpha}|D_{n1k}(t,x^\alpha) - ED_{n1k}(t,x^\alpha)| = o(\log n/(nh_{n1})^{1/2}) \qquad (36)
$$

almost surely as $n \to \infty$. In addition,

$$
\begin{aligned}
&ED_{n1k}(t,x^\alpha) \\
&= \frac{1}{h_{n1}}\int B_k(t,x^\alpha,\xi^{(-\alpha)})S_\alpha(t,x^\alpha,\xi^{(-\alpha)})K_1\left(\frac{x^\alpha-\xi^\alpha}{h_1}\right) \\
&\qquad\qquad\qquad \times p(t,\xi^\alpha,\xi^{(-\alpha)})dtd\xi^\alpha,d\xi^{(-\alpha)} \\
&= \int B_k(t,x^\alpha,X^{(-\alpha)})S_\alpha(t,x^\alpha,X^{(-\alpha)})k_1(v)P(t,x^\alpha-h_{n1}v,X^{(-\alpha)})dtdvdX^{(-\alpha)} \\
&= D_{1k}(t,x^\alpha) + O(h_{n1}^r) \qquad (37)
\end{aligned}
$$

uniformly over $(t,x^\alpha) \in S_T \times S_X$. The lemma follows by combining Equation (35), (36), and (37). ∎

Define

$$\begin{aligned}
\hat{D}_{n2}(t,x^\alpha) &= (nh_{n1})^{-1}\sum_{i=1}^{n} B_n(t,x^\alpha,X_i^{(-\alpha)})S_n(t|x^\alpha,X_i^{(-\alpha)}) \\
&\qquad\qquad \times p_{n\alpha}(x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \\
D_{n2}(t,x^\alpha) &= (nh_{n1})^{-1}\sum_{i=1}^{n} B(t,x^\alpha,X_i^{(-\alpha)})S(t|x^\alpha,X_i^{(-\alpha)}) \\
&\qquad\qquad \times p_{\alpha}(x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \\
D_2(t,x^\alpha) &= \int B(t,x^\alpha,X^{(-\alpha)})S(t|x^\alpha,X^{(-\alpha)})p_\alpha(x^\alpha,X^{(-\alpha)}) \\
&\qquad\qquad \times p(t,x^\alpha,X^{(-\alpha)})dtdX^{(-\alpha)}
\end{aligned}$$

The k-th component $\hat{D}_{n2k}(t,x^\alpha)$ and $D_{n2k}(t,x^\alpha)$ of $\hat{D}_{n2}(t,x^\alpha)$ and $D_{n2}(t,x^\alpha)$ follow by

$$\begin{aligned}
\hat{D}_{n2k}(t,x^\alpha) &= (nh_{n1})^{-1}\sum_{i=1}^{n} B_{nk}(t,x^\alpha,X_i^{(-\alpha)})S_n(t|x^\alpha,X_i^{(-\alpha)}) \\
&\qquad\qquad \times p_{n\alpha}(x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \\
D_{n2k}(t,x^\alpha) &= (nh_{n1})^{-1}\sum_{i=1}^{n} B_k(t,x^\alpha,X_i^{(-\alpha)})S(t|x^\alpha,X_i^{(-\alpha)}) \\
&\qquad\qquad \times p_{\alpha}(x^\alpha,X_i^{(-\alpha)})K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \\
D_{2k}(t,x^\alpha) &= \int B_k(t,x^\alpha,X^{(-\alpha)})S(t|x^\alpha,X^{(-\alpha)})p_\alpha(x^\alpha,X^{(-\alpha)}) \\
&\qquad\qquad \times p(t,x^\alpha,X^{(-\alpha)})dtdX^{(-\alpha)}
\end{aligned}$$

Lemma 7 *Under the assumption 1-4,*

$$\hat{D}_{n2}(t,x^\alpha) = D_2(t,x^\alpha) + O(h_{n1}^r) + o(\log n/(nh_{n1})^{1/2})$$

almost surely uniformly over $(t,x^\alpha) \in S_T \times S_X$.

Proof (Part A) By linear approximation of $B_{nk}S_n p_{n\alpha} - B_k S p_\alpha$ and triangle inequality,

$$
\begin{aligned}
&|B_{nk}(t, x^\alpha, X_i^{(-\alpha)})S_n(t|x^\alpha, X_i^{(-\alpha)})P_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) \\
&\qquad\qquad -B_k(t, x^\alpha, X_i^{(-\alpha)})S(t|x^\alpha, X_i^{(-\alpha)})P_\alpha(x^\alpha, X_i^{(-\alpha)})| \\
&\leq |B_{nk}(t, x^\alpha, X_i^{(-\alpha)}) - B_k(t, x^\alpha, X_i^{(-\alpha)})||S(t|x^\alpha, X_i^{(-\alpha)})P_\alpha(x^\alpha, X_i^{(-\alpha)})| \\
&+|S_n(t|x^\alpha, X_i^{(-\alpha)}) - S(t|x^\alpha, X_i^{(-\alpha)})||B_k(t, x^\alpha, X_i^{(-\alpha)})P_\alpha(x^\alpha, X_i^{(-\alpha)})| \\
&+|P_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) - P_\alpha(x^\alpha, X_i^{(-\alpha)})|B_k(t, x^\alpha, X_i^{(-\alpha)})S(t|x^\alpha, X_i^{(-\alpha)})| \\
&+R_n^2(t, x^\alpha, X_i^{(-\alpha)})
\end{aligned} \tag{38}
$$

where the remainder term $R_n^2(t, x^\alpha, X_i^{(-\alpha)}) = O[\{|B_{nk} - B_k||S_n - S|\}|P_\alpha| + \{|S_n - S||P_{n\alpha} - P_\alpha|\}|B_k| + \{|B_{nk} - B_k||P_{n\alpha} - P_\alpha|\}|S| + |B_{nk} - B_k||S_n - S||P_{n\alpha} - P_\alpha|]$. In addition,

$$
\begin{aligned}
|S_n(t|x^\alpha, X_i^{(-\alpha)}) \quad - \quad & S(t|x^\alpha, X_i^{(-\alpha)})| \\
&= |\frac{S_{n1}(t, x^\alpha, X_i^{(-\alpha)})}{P_n(x^\alpha, X_i^{(-\alpha)})} - \frac{S_1(t, x^\alpha, X_i^{(-\alpha)})}{P(x^\alpha, X_i^{(-\alpha)})}| \\
&\leq \frac{1}{P(x^\alpha, X_i^{(-\alpha)})}|S_{n1}(t, x^\alpha, X_i^{(-\alpha)}) - S_1(t, x^\alpha, X_i^{(-\alpha)})| \\
&+\frac{S_1(t, x^\alpha, X_i^{(-\alpha)})}{P^2(x^\alpha, X_i^{(-\alpha)})}|P_n(x^\alpha, X_i^{(-\alpha)}) - P(x^\alpha, X_i^{(-\alpha)})| \\
&+R_n^S(t, x^\alpha, X_i^{(-\alpha)})
\end{aligned} \tag{39}
$$

where the remainder term $R_n^s(t, x^\alpha, X_i^{(-\alpha)}) = O[(P_n - P)^2 + |(P_n - P)(S_{n1} - S_1)|]$. Combining Equation (38) and (39) with assumption 4 and applying Equation (33) and (34) in Lemma 6 gives

$$
\begin{aligned}
&|\hat{D}_{n2k}(t, x^\alpha) - D_{n2k}(t, x^\alpha)| \\
&\leq \sup|B_{nk}(t, x^\alpha, X_i^{(-\alpha)})S_n(t|x^\alpha, X_i^{(-\alpha)})P_{n\alpha}(x^\alpha, X_i^{(-\alpha)}) \\
&\qquad\qquad -B_k(t, x^\alpha, X_i^{(-\alpha)})S(x^\alpha, X_i^{(-\alpha)})P_\alpha(x^\alpha, X_i^{(-\alpha)})| \\
&\quad \times(nh_1)^{-1}\sum_{i=1}^{n}|K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right)| \\
&= o_p(1)O(1) \\
&= o_p(1)
\end{aligned} \tag{40}
$$

uniformly over $(t, x^\alpha) \in S_T \times S_X$.

(Part B) As in Lemma 1, $K_1\left(\frac{x^\alpha - \xi^\alpha}{h_{n1}}\right)$ is Euclidean, and $B_k(t, x^\alpha, X^{(-\alpha)})$ and $S(t|x^\alpha, X^{(-\alpha)})$, and $P_\alpha(x^\alpha, X^{(-\alpha)})$ are Euclidean because these are single functions. Hence, by Lemma 2.14 of Pakes and Pollard (1989), the summand of D_{n2} is Euclidean. Applying Theorem

2.37 of Pollard (1984) yields,

$$\sup_{t,x^\alpha} |D_{n2k}(t,x^\alpha) - ED_{n2k}(t,x^\alpha)| = o(\log n/(nh_{n1})^{1/2}) \tag{41}$$

almost surely as $n \to \infty$. In addition,

$$\begin{aligned} ED_{n2k}(t,x^\alpha) &= \frac{1}{h_{n1}} \int B(t,x^\alpha,\xi^{(-\alpha)}) S(t|x^\alpha,\xi^{(-\alpha)}) p_\alpha(t,x^\alpha,\xi^{(-\alpha)}) \\ &\qquad \times K_1(\frac{x^\alpha - \xi^\alpha}{h_{n1}}) p(t,\xi^\alpha,\xi^{(-\alpha)}) dt d\xi^\alpha d\xi^{(-\alpha)} \\ &= \int B(t,x^\alpha,X^{(-\alpha)}) S(t|x^\alpha,X^{(-\alpha)}) p_\alpha(x^\alpha,X^{(-\alpha)}) \\ &\qquad \times K_1(v) p(t,x^\alpha - h_{n1}v, X^{(-\alpha)}) dt dv dX^{(-\alpha)} \\ &= D_{2k}(t,x^\alpha) + O(h_{n1}^r) \end{aligned} \tag{42}$$

The Lemma follows by combining Equation (40)-(42). ∎

Define

$$\begin{aligned} \hat{G}_n(t,x^\alpha) &= \frac{1}{nh_1} \sum_{i=1}^n B_n(t,x^\alpha,X_i^{(-\alpha)}) B_n(t,x^\alpha,X_i^{(-\alpha)})' K_1(\frac{x^\alpha - X_i^\alpha}{h_1}) \\ G_n(t,x^\alpha) &= \frac{1}{nh_1} \sum_{i=1}^n B(t,x^\alpha,X_i^{(-\alpha)}) B(t,x^\alpha,X_i^{(-\alpha)})' K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right) \\ G(t,x^\alpha) &= \int B(t,x^\alpha,X^{(-\alpha)}) B(t,x^\alpha,X^{(-\alpha)})' p(t,x^\alpha,X^{(-\alpha)}) dt dX^{(-\alpha)} \end{aligned}$$

Lemma 8 *Under the assumption 1-4,*

$$\hat{G}(t,x^\alpha) = G(t,x^\alpha) + O(h_{n1}^r) + o(\log n/(nh_{n1})^{-1/2})$$

almost surely uniformly over $(t,x^\alpha) \in S_T \times S_X$

Proof (Part A) Note that $B(t,x^\alpha,X_i^{(-\alpha)})$ is bounded in compact support $S_T \times S_X$, then the linear approximation of $B_n B_n' - BB'$ gives

$$\begin{aligned} &|B_n(t,x^\alpha,X_i^{(-\alpha)}) B_n(t,x^\alpha,X_i^{(-\alpha)})' \\ &\qquad\qquad - B(t,x^\alpha,X_i^{(-\alpha)}) B(t,x^\alpha,X_i^{(-\alpha)})'| \\ &\leq B(t,x^\alpha,X_i^{(-\alpha)}) |B_n(t,x^\alpha,X_i^{(-\alpha)}) - B(t,x^\alpha,X_i^{(-\alpha)})|' \\ &\quad + |B_n(t,x^\alpha,X_i^{(-\alpha)}) - B(t,x^\alpha,X_i^{(-\alpha)})| B(t,x^\alpha,X_i^{(-\alpha)})' \\ &\quad + R_n^B(t,x^\alpha,X_i^{(-\alpha)}) \end{aligned} \tag{43}$$

where the remainder term $R_n^B(t,x^\alpha,X_i^{(-\alpha)}) = O[|(B_n - B)(B_n - B)'|]$. From Lemma 6, we know that and sup $|B_n(t,x^\alpha,X_i^{(-\alpha)}) - B(t,x^\alpha,X_i^{(-\alpha)})| = o_p(1)$. Therefore,

$$
\begin{aligned}
&|\hat{G}_n(t,x^\alpha) - G_n(t,x^\alpha)| \\
&\le \sup|B_n(t,x^\alpha,X_i^{(-\alpha)})B_n(t,x^\alpha,X_i^{(-\alpha)})' \\
&\qquad\qquad\qquad -B(t,x^\alpha,X_i^{(-\alpha)})B(t,x^\alpha,X_i^{(-\alpha)})'| \\
&\quad \times \frac{1}{nh_1}\sum_{i=1}^{n}|K_1\left(\frac{x^\alpha - X_i^\alpha}{h_{n1}}\right)| \\
&= o_p(1)O(1) \\
&= o_p(1) \qquad (44)
\end{aligned}
$$

(Part B) As in Lemma 1, $K_1\left(\frac{x^\alpha-\xi^\alpha}{h_1}\right)$ is Euclidean, and $B_k(t,x^\alpha,X^{(-\alpha)})$ are Euclidean for each k because these are single functions. Hence, by Lemma 2.14 of Pakes and Pollard (1989), the summand of G_n is Euclidean. Applying Theorem 2.37 of Pollard (1984) yields,

$$\sup_{t,x^\alpha}|G_n(t,x^\alpha) - EG_n(t,x^\alpha)| = o(\log n/(nh_1)^{1/2}) \qquad (45)$$

almost surely as $n \to \infty$. In addition,

$$
\begin{aligned}
EG_n(t,x^\alpha) &= \frac{1}{h_1}\int B(t,x^\alpha,\xi^{(-\alpha)})B(t,x^\alpha,\xi^{(-\alpha)})'K_1(\frac{x^\alpha-\xi^\alpha}{h_1}) \\
&\qquad \times p(t,\xi^\alpha,\xi^{(-\alpha)})dtd\xi^\alpha d\xi^{(-\alpha)} \\
&= \int B(t,x^\alpha,X^{(-\alpha)})B(t,x^\alpha,X^{(-\alpha)})'K_1(v) \\
&\qquad \times p(t,x^\alpha - h_1 v,X^{(-\alpha)})dtdvdX^{(-\alpha)} \\
&= G(t,x^\alpha) + O(h_1^r) \qquad (46)
\end{aligned}
$$

The lemma follows by Equation (44)-(46). ∎

Define $\hat{D}_n = \hat{D}_{n1} - \hat{D}_{n2}$ and $D = D_1 - D_2$. We know that by Lemma 6 and Lemma 7, $\hat{D} = D + O(h_{n1}^r) + o(\log n/(nh_{n1})^{1/2})$.

Lemma 9

$$M_n^\alpha(x^\alpha) = M_\alpha(x^\alpha) + O(h_{n1}^r) + o(\log n/(nh_{n1})^{1/2}) \qquad (47)$$

almost surely uniformly over $x^\alpha \in S_X$

Proof

$$
\begin{aligned}
& M_n^\alpha(x^\alpha) - M^\alpha(x^\alpha) \\
& = \int_{x_0^\alpha}^{x^\alpha} dx^\alpha \int_{S_T} dt(\hat{G}_n^{-1}(t, x^\alpha)\hat{D}_n^(t, x^\alpha) - G^{-1}(t, x^\alpha)D(t, x^\alpha)) \\
& = \int dx^\alpha \int dt G^{-1}(t, x^\alpha)\hat{D}_n(t, x^\alpha) \\
& - \int dx^\alpha \int dt G^{-1}(t, x^\alpha)\hat{G}_n(t, x^\alpha)G^{-1}(t, x^\alpha)D(t, x^\alpha) + o_p(1) \\
& = \int dx^\alpha \int dt G^{-1}(t, x^\alpha)D(t, x^\alpha) \\
& - \int dx^\alpha \int dt G^{-1}(t, x^\alpha)\hat{G}_n(t, x^\alpha)G^{-1}(t, x^\alpha)D(t, x^\alpha) \\
& + O(h_{n1}^r) + o(\log n/(nh_{n1})^{1/2}) \\
& = O(h_{n1}^r) + o(\log n/(nh_{n1})^{1/2}) \\
& = o_p(1)
\end{aligned}
$$

by the result of Lemma 6 to Lemma 8.

국문초록

오 나 래
경제학부 경제학전공
서울대학교 대학원

이 논문에서는 잠재 실패 시간이 임의의 함수 형태로 변환 되어있는 경우에 다중 위험 모델의 식별과 추정 방법에 대해서 다룬다. 우리는 잠재 위험 실패 시간이 비모수 가산 회귀 모형에 의해 생성된다고 가정하였다. 이 논문에서 다루는 모형은 로그 선형 모형, 결합 비례 위험 모형, 가속된 파괴 시점 모형, 그리고 선형 변환 모형 등을 포함하는 일반화된 모형이다. 임의의 가산 함수의 식별은 '부분적분' 방법을 사용하여 가능함이 확인되었다. 주어진 식별 결과에 기반하여 균등 일치 추정량을 제안하였다.

주요어 : 다중 위험 모델, 변환 모델, 식별, 부분적분, 추정
학 번 : 2011-20178